第1章　確率分布とその基本的な計算

abstract

1　Introduction

確率分布は、事象の起こりやすさを表現する数学の手法です。統計学では、店の来客数や顧客の滞在時間など、具体的な分析対象をモデル化して把握するために確率分布が応用されています。また、さまざまな統計学の手法の仕組みや性質を理解する際にも確率分布の知識は有用です。

確率分布を使いこなすためには、典型的な確率分布を抑えること、確率分布の特徴を数値に要約する方法を知り、計算できるようになることが大切です。そこで今回は、典型的な確率分布として、二項分布・Poisson（ポアソン）分布・ガンマ分布・ベータ分布を紹介します。また、数値要約の方法として期待値と分散を紹介します。

2　確率分布とはなにか

本節では、まず確率変数を導入し、確率分布の基本的なイメージを説明します。その後、確率変数を離散型と連続型の二つに分類し、それぞれの場合で確率分布をより正確に定義します。

2.1　確率変数と確率分布

確率変数 (random variable) は、ランダムな試行の結果によってその値が決まるような変数のことです。例えば、一日の店の来客数はその日によって値が異なります。ある日は 5 名、また別の日は誰も来ないかもしれません。このように、一日の店の来客数は試行によって値が変化するため確率変数です。

確率分布 (probability distribution) は、確率変数の取りうる値とその確率の対応関係を表すことで、さまざまな事象の発生確率を計算できるようにしたものです。確率変数として、公平なコインを2回トスしたときに表が出る回数を考えます。この確率変数の確率分布は、以下の表で表されます。

表が出る回数	**0**	**1**	**2**
確率	1/4	1/2	1/4

Table 1　確率分布の例

この表からはさまざまな事象が起こる確率を計算できます。例えば1回以上表が出るような事象が起こる確率は、表が1回出る確率 $(1/2)$ と表が2回出る確率 $(1/4)$ を合計して $1/2 + 1/4 = 3/4$ だとわかります。

確率分布が持つ「さまざまな事象が起こる確率を計算できる」という性質は実務上でも重要な役割を果たします。一日の店の来客数の確率分布からは「30 名以上のお客さんが来る確率は 10% 以下」などの情報を得ることができます。これは適切な人員配置の計画に役立ちます。

ロボットアームの異常を振動センサーの波形から検知する課題でも確率分布は有用です。正常時の波形のスパイク数の確率分布がわかれば、「100 回以上スパイクが含まれる確率は 5% 以下」といった情報を引き出すことができます。これは異常を予測するためのルールを作るのに役立ちます。

Remark　確率変数や確率分布を数学的に定義することは簡単ではないため、ここでは直感的な説明に留めました。興味のある方は、測度論的確率論を勉強してみてください。■

2.2　離散型確率変数と連続型確率変数

飛び飛びの値をとるような確率変数を**離散型確率変数**といいます。例えば、コイントスで表が出た回数、一日の店の来客数のように数え上げられるようなものが離散型確率変数です。一方で連続的な値をとるような確率変数を**連続型確率変数**といいます。例えば、顧客の滞在時間やweb広告のクリック率は連続型確率変数です。

次に進む前に、確率変数を離散型と連続型に区別することが重要な理由を説明します。以降の節では、2.1節より正確に確率分布を定義します。そしてこのとき、私たちは離散型確率変数と連続型確率変数それぞれの場合で確率分布を定義していくのです。「なぜ分けて定義すると都合が良いのか？」その理由も含めて解説していきましょう。

2.3　離散型確率変数の確率分布

離散型確率変数の確率分布は、確率変数 X が値 x をとる確率の表を作ることで定義されます。数学的には、表のかわりに以下のような関数を考えます。

$$f_X(x) = \mathbb{P}[X = x]$$

右辺 $\mathbb{P}[X = x]$ は確率変数 X が値 x をとる確率を意味する数式です。

この関数 $f_X(x)$ のことを**確率関数** (probability mass function, pmf) といいます。確率関数を使えば、さまざまな事象の発生確率を計算できます。つまり、離散型確率変数の確率分布は、確率関数の具体的な数式を与えることで定義されます。

例題　確率 p で表が出るコインをトスする試行を表が出るまで繰り返します。各試行は独立であるとします。表が出るまでコイントスした回数を確率変数 X と表すとき、X の確率関数を求めてください。

解答　確率変数 X は値 $x = 1, 2, 3, \cdots$ と正の整数をとることができます。特に $X = x$ になるときは、裏が $x-1$ 回出た後に表が 1 回出る事象に対応しています。図で表すなら以下の事象です。

$$\underbrace{\text{裏}, \text{裏}, \cdots, \text{裏}}_{x-1\text{ 回}}, \text{表}$$

裏が出る確率は1回あたり $1-p$、表が出る確率は1回あたり p なので確率関数は

$$\mathbb{P}[X = x] = p(1-p)^{x-1}, \qquad x = 1, 2, 3, \cdots$$

だとわかります。■

Remark　このような確率関数で定義される確率分布を**幾何分布** (geometric distribution) といいます。幾何分布には二通りの定義があり「表が出るまでに裏が出た回数」を確率変数 X に設定した場合もあるので注意が必要です。この場合の確率関数は
$f(x) = p(1-p)^x$ になります。■

2.4　連続型確率変数の確率分布

離散型確率変数では確率関数を用いて確率分布を定義しました。ところが連続型確率変数の場合、この方法は適用できません。なぜなら、特殊な場合を除いて確率関数の値が 0 になるからです。例えば、顧客の滞在時間が正確に 10 分ぴったりである確率は一般に 0 です。このため確率関数を用いても、さまざまな事象の確率を計算することができず、結果的に確率分布を定義することができません。

この問題を解決し、確率分布を定義できるようにするための工夫としては「値の範囲」に注目することが考えられます。例えば、顧客が 10 分から 20 分の間店に滞在する確率は一般に 0 ではありません。そこで、次のような関数 $f(x)$ を考えてみましょう。

$$\mathbb{P}[a \leq x \leq b] = \int_a^b f(x)dx$$

つまり $a \leq x \leq b$ の範囲で関数 $f(x)$ のグラフの面積を計算すると、確率変数 X の値が a 以上 b 以下の範囲に収まる確率を求められるような関数です。これを**確率密度関数** (probability density function, pdf) といいます。

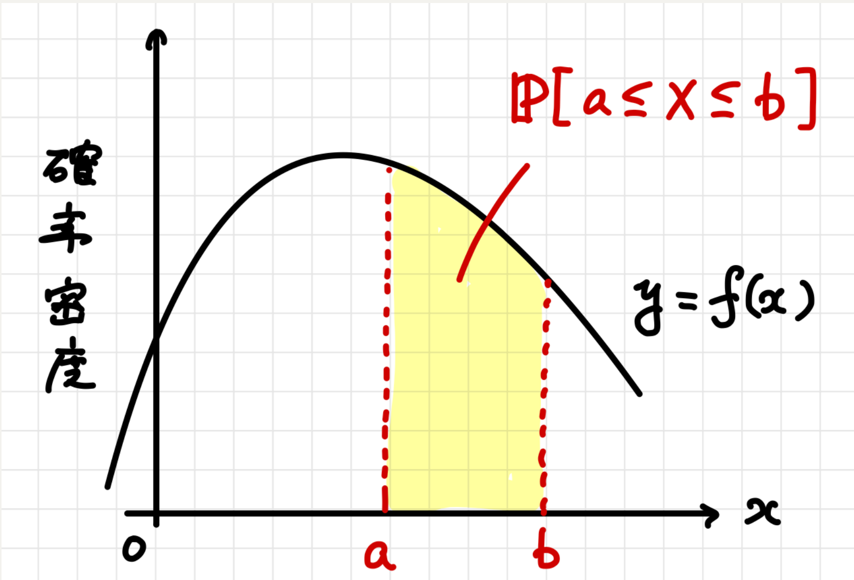

Figure 1　確率密度関数のイメージ

確率密度関数を見つけることができれば様々な事象の発生確率を計算することができます。つまり連続型確率変数の確率分布は、確率密度関数の具体的な数式を与えることで定義されます。

Remark　すべての連続型確率変数に対して確率密度関数が必ず存在するわけではありませんが、確率密度関数が存在しない場合を議論することは一般的ではありません。多くの実用的な場面では、確率密度関数は連続型確率変数の確率分布を定義するための有効な手段を与えています。■

例題　c を正の定数とします。確率変数 X の確率分布が以下のような確率密度関数で定義されるとします。

$$f_X(x) = \begin{cases} cx^2, & -1 \leq x \leq 1 \\ 0, & \text{その他} \end{cases}$$

以下の問に答えてください。

[1]　c の値を求めてください。

[2]　確率密度関数のグラフをかいてください。

[3]　確率変数 X が $1/2$ 以上の値をとる確率 $\mathbb{P}[X \geq 1/2]$ の値を計算してください。

解答

[1]　$\mathbb{P}[-1 \leq X \leq 1] = 1$ が成り立つことに注意します。確率密度関数を用いて左辺を計算すると、以下のように c を用いた式で表すことができます。

$$\mathbb{P}[-1 \leq X \leq 1] = c\int_{-1}^{1} x^2 dx = c\left[\frac{1}{3}x^3\right]_{-1}^{1} = \frac{2}{3}c$$

これで $\dfrac{2}{3}c = 1$ が得られたので、$c = \dfrac{3}{2}$ だとわかります。

[2]　グラフは以下のようになります。

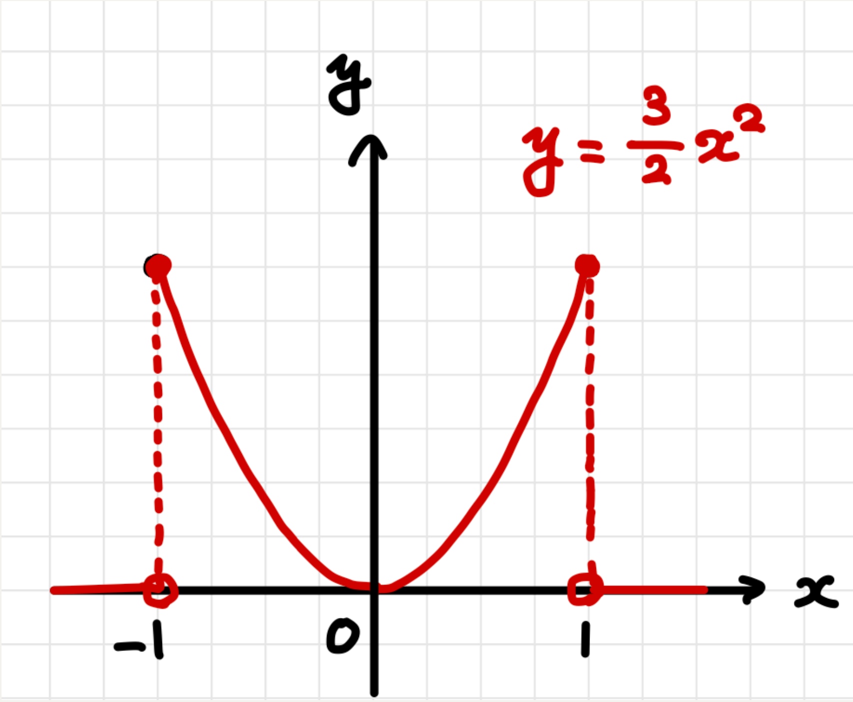

Figure 2　[2]の解答

[3]　確率密度関数の定義を用いれば、

$$\mathbb{P}\left[X \geq \frac{1}{2}\right] = \frac{3}{2}\int_{1/2}^{1} x^2 dx = \frac{3}{2}\left[\frac{1}{3}x^3\right]_{1/2}^{1} = \frac{7}{16}$$

と計算できます。なお、この確率の計算をグラフで表すと

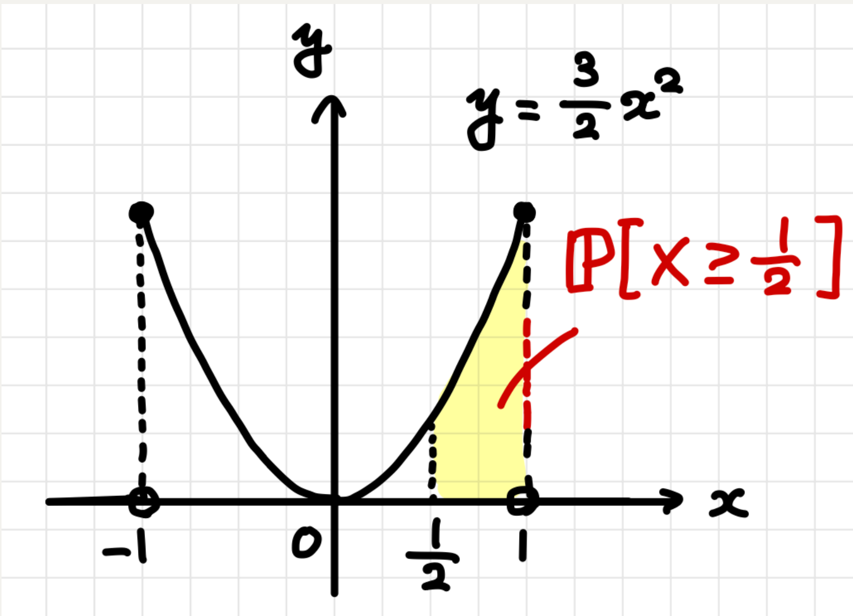

Figure 3　計算したい面積

の黄色の部分の面積を求めることに対応しています。■

Remark　このような確率密度関数で定義される確率分布を**標準U-二次分布** (standard U-quadratic distribution) といいます。計算練習にはちょうど良いですが、実務で使われることは少ないように思います。■

3　期待値と分散

本節では、期待値と分散を定義します。

3.1　期待値

期待値 (expected value) は、確率変数がとりうる各値をそれらの確率で重みづけした平均値として定義されます。例えば、1回のコイントスで表が出る回数 X について考えます。表が出る確率を p とすると、期待値 $\mathbb{E}[X]$ は次のように計算されます。

$$\mathbb{E}[X] = 0 \times (1-p) + 1 \times p = p$$

公平なコイン ($p = 1/2$) の場合、期待値は $1/2$ になります。一方、表が出やすいコイン ($p > 1/2$) の場合、期待値も $1/2$ より大きくなります。

期待値は確率分布の特性を表す重要な数値要約の一つです。特にコイントスの例からもわかるように、確率変数が平均的にとる値や、確率関数や確率密度関数のグラフの中心的な位置を知ることができます。

確率変数 X が離散型の場合には、確率関数 $f_X(x) = \mathbb{P}[X]$ を用いて次のように定義されます。

$$\mathbb{E}[X] = \sum_{x:X\text{がとりうる値}} x f_X(x)$$

一方で連続型の場合には、確率関数のかわりに確率密度関数を用いて定義します。

$$\mathbb{E}[X] = \int_{X\text{がとりうる値の範囲}} x f_X(x) dx$$

例題　次の確率密度関数で定義される確率分布に対して、期待値を計算してください。

$$f_X(x) = \begin{cases} \dfrac{3}{2}x^2, & -1 \leq x \leq 1 \\ 0, & \text{その他} \end{cases}$$

解答　この確率分布に従っている確率変数を X と表すことにします。期待値 $\mathbb{E}[X]$ は以下のように計算できます。

$$\begin{aligned}\mathbb{E}[X] &= \int_{-1}^{1} x \times \frac{3}{2}x^2 dx \\ &= \frac{3}{2}\int_{-1}^{1} x^3 dx \\ &= \frac{3}{2}\left[\frac{1}{4}x^4\right]_{-1}^{1} \\ &= 0\end{aligned}$$

従って期待値は $\mathbb{E}[X] = 0$ です。■

Remark　この例題の解答として、確率密度関数のグラフが $x = 0$ で対称になっていることから $\mathbb{E}[X] = 0$ と結論づけるという直感的な方法もあります。ただしこの方法は注意が必要で、$xf_X(x) = \dfrac{3}{2}x^3$ が積分できるという事実を暗黙に使っています。統計検定の解答を書く際には、言及したほうが良いと思います。

この話には、コーシー分布という実務上も重要な確率分布が背景にあります。コーシー分布は、$x = 0$ で対称なグラフを持つ確率密度関数で定義されるにも関わらず、$xf_X(x)$ が積分できない関数になるため期待値は存在しません。もちろん、$\mathbb{E}[X] = 0$ ではありません。これにはコーシー分布の確率密度関数の裾の重さが関連していますが、ここでは詳しく解説しません。■

3.2　一般の場合の期待値

ここでは確率変数 X の式で表される確率変数 $g(X)$ に対する期待値 $\mathbb{E}[g(X)]$ を考えます。例えば、公平なコインを2回トスして、表が出るごとに100円貰えるようなゲームを考えます。このときコインの表が出る回数を X と表すなら、このゲームを通して貰えるお金の総額 $g(X)$ は

$$g(X) = 100X$$

と表すことができます。今回興味があるのは、このゲームを通して貰えるお金 $g(X) = 100X$ の期待値 $\mathbb{E}[100X]$ です。

この期待値は次のように計算できることがわかります。証明はAppendixを参照してください。もし証明を理解するのが難しくても、一般の場合の期待値の定義だと思って、結果を認めたまま先に進んで構いません。

定理　確率変数 X を離散型とし、この確率関数を $f_X(x)$ と表すことにします。このとき、$g(X)$ の期待値は以下のように計算できます。

$$\mathbb{E}[g(X)] = \sum_{x:X\text{の取りうる値}} g(x) f_X(x)$$

連続型の場合も同様に $\mathbb{E}[g(X)] = \int_{X\text{がとりうる値の範囲}} g(x) f_X(x) dx$ と計算できます。

この定理を用いて、先ほどのゲームを通して貰えるお金の総額 $g(X) = 100X$ の期待値を計算してみましょう。X の確率分布には2.1節に与えられているものを用います。

$$\mathbb{E}[100X] = 0 \times \frac{1}{4} + 100 \times \frac{1}{2} + 200 \times \frac{1}{4} = 100$$

以上で、ゲームを通して貰えるお金の総額の期待値は 100 円だとわかりました。

例題　確率変数 X が以下の確率密度関数で定義される確率分布に従っているとします。X^2 の期待値 $\mathbb{E}[X^2]$ を計算してください。

$$f_X(x) = \begin{cases} \frac{3}{2}x^2, & -1 \leq x \leq 1 \\ 0, & \text{その他} \end{cases}$$

解答　期待値 $\mathbb{E}[X^2]$ は以下のように計算できます。

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-1}^{1} x^2 \times \frac{3}{2}x^2 dx \\ &= \frac{3}{2}\int_{-1}^{1} x^5 dx \\ &= \frac{3}{2}\left[\frac{1}{5}x^5\right]_{-1}^{1} \\ &= \frac{3}{5} \end{aligned}$$

従って、X^2 の期待値は $\mathbb{E}[X^2] = 3/5$ です。■

3.3　期待値の線形性

一般の場合の期待値を計算するとき、以下の公式を知っておくと便利です。これを**期待値の線形性**といいます。

定理　X を確率変数、c を実数とします。以下の公式が成り立ちます。
(1)　$\mathbb{E}[g(X)+h(X)]=\mathbb{E}[g(X)]+\mathbb{E}[h(X)]$
(2)　$\mathbb{E}[cg(X)]=c\mathbb{E}[g(X)]$

証明　連続型の場合で証明を与えます。離散型の場合も同様に証明できるので、興味がある方はぜひ考えてみてください。なお、以下では X の確率密度関数を $f_X(x)$、X が取りうる値の範囲を $\mathcal{X}$ と表すことにします。

(1)　期待値の定義と積分の性質を用いて証明することができます。

$$\begin{aligned}\mathbb{E}[g(X)+h(X)]&=\int_{\mathcal{X}}(g(x)+h(x))f_X(x)dx\\&=\int_{\mathcal{X}}g(x)f_X(x)dx+\int_{\mathcal{X}}h(x)f_X(x)dx\\&=\mathbb{E}[g(X)]+\mathbb{E}[h(X)]\end{aligned}$$

(2)　こちらも(1)と同様に、期待値の定義と積分の性質を用いて証明することができます。

$$\begin{aligned}\mathbb{E}[cg(X)]&=\int_{\mathcal{X}}cg(x)f_X(x)dx\\&=c\int_{\mathcal{X}}g(x)f_X(x)dx\\&=c\mathbb{E}[g(X)]\end{aligned}$$

以上で(1)、(2)ともに証明できました。■

Remark　(1)には、以下のような一般化が成り立ちます。

$$\mathbb{E}[X+Y]=\mathbb{E}[X]+\mathbb{E}[Y]$$

この事実を証明するには多次元確率分布を導入する必要があるため、第3章で改めて解説します。■

3.4　分散

分散（variance）は、確率変数 X の値の散らばり具合を表す統計量です。具体的には、次のように考えて定義します。

1. 確率変数 X の**偏差**（derivation）を確率変数 X と期待値 $\mathbb{E}[X]$ の差 $X - \mathbb{E}[X]$ と定義する。
2. 偏差の大きさの平均的な値が確率変数 X の値の散らばり具合を表現すると考えられることに注目する。
3. 平均をとるとき、偏差のままだと正の数と負の数で打ち消しあってしまうため、偏差の二乗 $(X - \mathbb{E}[X])^2$ をとって平均する。つまり、

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

   を考える。これを確率変数 X の分散という。

分散は大きさが二乗になったままで解釈が難しいため、確率変数のもともとの値と大きさを揃えるために分散のルートをとったもの $\sqrt{\mathbb{V}[X]}$ を確率変数 X の**標準偏差**（standard deviation）という。

例題　次の確率密度関数で定義される確率分布に対して、分散を計算してください。

$$f_X(x) = \begin{cases} \frac{3}{2}x^2, & -1 \leq x \leq 1 \\ 0, & \text{その他} \end{cases}$$

ただし、期待値が 0 になることは用いてよいものとします。

解答　この確率分布に従っている確率変数を X と表すことにします。分散 $\mathbb{V}[X]$ は以下のように計算できます。

$$\begin{aligned} \mathbb{V}[X] &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \int_{-1}^{1} (x-0)^2 \times \frac{3}{2}x^2 dx \\ &= \frac{3}{2}\int_{-1}^{1} x^4 dx \\ &= \frac{3}{2}\left[\frac{1}{5}x^5\right]_{-1}^{1} \\ &= \frac{3}{5} \end{aligned}$$

従って分散は $\mathbb{V}[X] = 3/5$ です。■

Remark　偏差の二乗をとるかわりに、偏差の絶対値をとるという考え方もあります。この場合の散らばり具合を測る統計量は $\mathbb{E}[|X - \mathbb{E}[X]|]$ です。これを**平均絶対偏差**（mean absolute deviation, MAD）といいます。今後の勉強でわかることですが、分散は平均絶対偏差に比べて計算しやすい性質が多くあるため、最初に分散から勉強することが一般的です。■

3.5　分散の性質

分散には次のような性質が成り立ちます。特に(2)は**分散の公式**とよばれるものです。分散を計算するときは、定義通りに計算するより、分散の公式を用いたほうが楽に計算できることが多いので覚えておきましょう。

定理　X を確率変数、c を実数とします。以下の公式が成り立ちます。
(1)　$\mathbb{V}[cX] = c^2\mathbb{V}[X]$
(2)　$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

証明　3.3節で導入した期待値の性質を応用して証明します。

(1)　分散を定義に従って期待値で表現すると、3.3節で導入した期待値の性質 $\mathbb{E}[cg(X)] = c\mathbb{E}[g(X)]$ を用いることで証明できます。

$$\begin{aligned}
\mathbb{V}[cX] &= \mathbb{E}[(cX - \mathbb{E}[cX])^2] \\
&= \mathbb{E}[(cX - c\mathbb{E}[X])^2] \\
&= \mathbb{E}[c^2(X - \mathbb{E}[X])^2] \\
&= c^2\mathbb{E}[(X - \mathbb{E}[X])^2] \\
&= c^2\mathbb{V}[X]
\end{aligned}$$

(2)　分散の定義に従って期待値で表現したあと、3.3節で導入した期待値の性質を用いて式を展開することで証明できます。ここで、期待値 $\mathbb{E}[X]$ は確率変数ではなく実数値として扱えるため、

$$\begin{aligned}
\mathbb{E}[X\mathbb{E}[X]] &= \mathbb{E}[X]\mathbb{E}[X] = \mathbb{E}[X]^2 \\
\mathbb{E}[\mathbb{E}[X]^2] &= \mathbb{E}[X]^2\mathbb{E}[1] = \mathbb{E}[X]^2
\end{aligned}$$

が成り立つことに注意しましょう。

$$\begin{aligned}\mathbb{V}[X] &= \mathbb{E}[(X-\mathbb{E}[X])^2] \\ &= \mathbb{E}[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

となり、分散の公式を証明することができました。

例題　確率変数 X が以下の確率密度関数で定義されているものとします。

$$f_X(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{その他} \end{cases}$$

以下の問に答えてください。

[1]　期待値 $\mathbb{E}[X]$ の値を求めてください。

[2]　分散 $\mathbb{V}[X]$ の値を分散の定義通りに計算してください。

[3]　分散 $\mathbb{V}[X]$ の値を分散の公式に従って計算してください。

解答

[1]　期待値は次のように計算できます。

$$\mathbb{E}[X] = \int_0^1 2x^2 dx = 2\int_0^1 x^2 dx = 2\left[\frac{1}{3}x^3\right]_0^1 = \frac{2}{3}$$

[2]　分散を定義通りに計算すると以下のようになります。

$$\begin{aligned}\mathbb{V}[X] &= \mathbb{E}[(X-\mathbb{E}[X])^2] \\ &= \int_0^1 2x\left(x-\frac{2}{3}\right)^2 dx \\ &= 2\int_0^1 x^3 dx - \frac{8}{3}\int_0^1 x^2 dx + \frac{8}{9}\int_0^1 x dx \\ &= 2\left[\frac{1}{4}x^4\right]_0^1 - \frac{8}{3}\left[\frac{1}{3}x^3\right]_0^1 + \frac{8}{9}\left[\frac{1}{2}x^2\right]_0^1 \\ &= \frac{1}{2} - \frac{8}{9} + \frac{4}{9} \\ &= \frac{1}{18}\end{aligned}$$

[3]　分散の公式に従って分散を計算すると以下のようになります。

$$\mathbb{E}[X^2] = \int_0^1 2x^3 dx = 2\int_0^1 x^3 dx = 2\left[\frac{1}{4}x^4\right]_0^1 = \frac{1}{2}$$

この結果を分散の公式に代入すると、

$$\begin{aligned}\mathbb{V}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \frac{1}{2} - \frac{4}{9} \\ &= \frac{1}{18}\end{aligned}$$

が得られます。[2]に比べて、計算が簡単になったことを実感できたでしょうか。■

Remark　定理(1) $\mathbb{V}[cX] = c^2\mathbb{V}[X]$ で係数が c^2 と二乗になる部分に違和感がある方がいるかもしれません。これは、分散の単位が確率変数のもともとの大きさの二乗になっていることに注目すれば、直感的に納得できるのではないでしょうか。■

Remark　定理(2)の証明について、$\mathbb{E}[X^2] = \mathbb{E}[X]^2$ は成り立たないのかという質問を受けますが、これは成り立ちません。 実際、$x = 0$ で左右対称な確率密度関数を持つ確率変数 X は、一般に期待値が $\mathbb{E}[X] = 0$ になります。このとき、期待値の二乗もまた $\mathbb{E}[X]^2 = 0$ です。しかし、確率変数の二乗 X^2 は 0 以上の値しかとらないため、期待値をとっても一般に $\mathbb{E}[X^2] \neq 0$ になります。■

Appendix　3.2節の定理の証明

定理　確率変数 X を離散型とし、この確率関数を $f_X(x)$ と表すことにします。このとき、$g(X)$ の期待値は以下のように計算できます。

$$\mathbb{E}[g(X)] = \sum_{x:X\text{の取りうる値}} g(x)f_X(x)$$

連続型の場合も同様に $\mathbb{E}[g(X)] = \int_{X\text{がとりうる値の範囲}} g(x)f_X(x)dx$ と計算できます。

証明　離散型の場合のみ証明します。確率変数 $Y = g(X)$ の期待値 $\mathbb{E}[g(X)]$ は、期待値の定義から以下のように計算されるはずです。

$$\mathbb{E}[g(X)] = \sum_{y:g(X)\text{のとりうる値}} y\mathbb{P}[g(X)=y]$$

総和記号のなかに現れる確率 $\mathbb{P}[g(X)=y]$ は、$g(x)=y$ を満たす値 x が起こる確率に等しいはずです。数式で表すなら以下のようになります。

$$\mathbb{P}[g(X)=y] = \sum_{x:g(x)=y} \mathbb{P}[X=x]$$

これを $\mathbb{E}[g(X)]$ の右辺に代入すると、以下のような式変形が得られます。

$$\begin{aligned}
\mathbb{E}[g(X)] &= \sum_{y:g(X)\text{のとりうる値}} y\mathbb{P}[g(X)=y] \\
&= \sum_{y:g(X)\text{のとりうる値}} y\left(\sum_{x:g(x)=y} \mathbb{P}[X=x]\right) \\
&= \sum_{y:g(X)\text{のとりうる値}} \left(\sum_{x:g(x)=y} y\mathbb{P}[X=x]\right) \\
&= \sum_{y:g(X)\text{のとりうる値}} \left(\sum_{x:g(x)=y} g(x)\mathbb{P}[X=x]\right) \\
&= \sum_{x:X\text{のとりうる値}} g(x)\mathbb{P}[X=x]
\end{aligned}$$

あとは確率関数の定義から $f_X(x)=\mathbb{P}[X=x]$ なので、確率変数 $g(X)$ の期待値が

$$\mathbb{E}[g(X)] = \sum_{x:X\text{のとりうる値}} g(x)f_X(x)$$

と計算できることがわかりました。■

第1章のまとめ

- 確率分布は、離散型確率変数の場合には確率関数、連続型確率変数の場合には確率密度関数の具体的な数式を与えることで定義される。
- 期待値は確率変数がとりうる値の平均値のことで、連続型確率変数 X の場合はその確率密度関数を $f_X(x)$ とすると、以下のように定義される。

$$\mathbb{E}[X] = \int_{Xがとりうる値の範囲} x f_X(x) dx$$

- 分散は確率変数が取りうる値の散らばり具合のことで、偏差の二乗の期待値 $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$ と定義される。
- 分散の値を計算するときは、定義から計算するよりも分散の公式 $\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ を用いて計算するほうが簡単になることが多い。

第1章の演習問題

問1　λ を正の実数とする。X を連続型確率変数とし、確率密度関数が

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x > 0) \\ 0 & (x \leq 0) \end{cases}$$

であるとする。以下の [1] から [3] に答えよ。

[1]　X の期待値 $\mathbb{E}[X]$ と分散 $\mathbb{V}[X]$ を求めよ。

[2]　n を正の整数とする。X^n の期待値 $\mathbb{E}[X^n]$ を求めよ 。

[3]　$Y = X^2$ とする。Y の期待値 $\mathbb{E}[Y]$ と分散 $\mathbb{V}[Y]$ を求めよ。

問2　p を $0 < p < 1$ を満たす実数とする。X を離散型確率変数とし、確率関数が

$$f(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, 3, \cdots \\ 0 & \text{その他} \end{cases}$$

であるとする。以下の [1] と [2] に答えよ。

[1]　X の期待値 $\mathbb{E}[X]$ を求めよ。

[2]　X の分散 $\mathbb{V}[X]$ を求めよ。